Title: Mechanisms and consequences of positive-strand RNA virus recombination

Authors: Kirsten Bentley* and David J. Evans

Affiliation: Biomedical Sciences Research Complex and School of Biology, University of St. Andrews

***Communicating author**: Dr. Kirsten Bentley, E: kb209@st-andrews.ac.uk, T: 01334 464801

Keywords: Recombination; Positive-strand RNA viruses; poliovirus; virus diversity; virus evolution.

Word count: 6383

Abbreviations: RdRp, RNA-dependent RNA polymerase; RF, recombinant form; CRE, *cis*-replicating element; BVDV, bovine viral diarrhoea virus; HCV, hepatitis C virus; WEE, western equine encephalitis virus; BMV, Brome mosaic virus; DI-RNA, defective-interfering RNA

Abstract

Genetic recombination in positive-strand RNA viruses is a significant evolutionary mechanism which drives the creation of viral diversity by formation of novel chimeric genomes. The process and its consequences, for example the generation of viruses with novel phenotypes, has been historically studied by analysis of the end products. More recently, with an appreciation that there are both replicative and non-replicative mechanisms at work, and with new approaches and techniques to analyse intermediate products, the viral and cellular factors that influence the process are becoming understood. The major influence on replicative recombination is the fidelity of viral polymerase, though RNA structures and sequence may also have an impact. In replicative recombination the viral polymerase is necessary and sufficient, though roles for other viral or cellular proteins may exist. In contrast, non-replicative recombination appears mediated solely by cellular components. Despite these insights, the relative importance of replicative and non-replicative mechanisms is not clear. Using single-stranded, positive sense, RNA viruses as exemplars we review the current state of understanding of the processes and consequences of recombination.

Introduction

RNA viruses are ubiquitous in nature as a consequence of their ability to evolve rapidly and to adapt to new environments. This rapid evolution, in turn, is partly dependent upon the high levels of genetic diversity that are a hallmark of RNA virus populations. This diversity primarily arises from the error prone nature of the viral RNA-dependent RNA polymerases (RdRp). During genome replication these introduce substitutions, insertions and deletions and typically exhibit error rates of 10^{-3} to 10^{-5} per nucleotide polymerized [1]. However, much larger scale variation can be introduced into the virus population through the analogous processes of reassortment in the segmented RNA viruses, and recombination, which can occur in both the segmented and non-segmented RNA virus families. To generate viable hybrid progeny, both reassortment and recombination require the co-infection of a single cell with compatible genomes. However, reassortment occurs as a result of the exchange of discrete segments during genome packaging into nascent particles ([2], and reviewed in [3]), while recombination results in the formation of genetic hybrids within the genomic segment by fundamentally different mechanistic processes.

A role for recombination in driving genetic diversity has been known for almost 100 years [4], but the ability of RNA viruses to exchange genetic material via recombination was

only discovered relatively recently. However, what is clear is that recombination is primarily a process of positive-strand RNA viruses and is observed very rarely in the negative-strand viruses (reviewed in [5]). For this reason, this review focuses on the single-stranded positive-sense RNA viruses. The first experimental evidence of recombination in these viruses arose from studies in poliovirus [6-8], but has since been shown to occur in a wide range of positive-sense RNA viruses [9-11], infecting all types of organisms from humans and animals, to plants and bacteria [12-14]. It seems likely that recombination is ubiquitous in the positive-strand RNA viruses.

Virus Evolution and the Consequences of Recombination

Because of the extensive genetic changes achievable through recombination, the process can result in rapid and extreme changes in virus phenotype, including escape from the immune response or antiviral therapy, changes in cell or host tropism and alterations of pathogenicity. Often, recombination can go unnoticed in a virus population, but where events are directly linked to novel outbreaks of disease these recombination events, and resulting viruses, are well documented. For example, all recipients of the live-attenuated trivalent oral poliovirus vaccine excrete type 1/type 3 recombinants without issue within a week of vaccination [15, 16]. However, in regions of low vaccine coverage, recombination between polio vaccine and co-circulating type C enteroviruses has been associated with vaccine-derived paralysis in vaccinees or their contacts. In the first such outbreak characterised, the virulent viruses responsible for paralysis of Haitian children were sequenced and shown to consist of the Sabin 1-derived 5' non-coding region (NCR) and capsid-coding region, recombined with co-circulating species C enteroviruses from which the majority of the non-structural proteins and 3' NCR were derived. At least four independent type 1 vaccine-derived recombinant poliovirus strains were identified in this outbreak, all of which also carried a 5' NCR mutation known to be associated with neurovirulence [17].

Subsequent prospective and retrospective studies [18-20] demonstrated how relatively frequently this type of intratypic recombination event in poliovirus could be observed. At the same time, epidemiological studies of circulating enterovirus species showed evidence for very extensive intratypic recombination [9, 21-23]. For the enterovirus serotypes (defined by the capsid-coding sequences to which antibodies are directed) tested this was characterised by the appearance, proliferation and subsequent disappearance of particular unique recombinant forms (RF) in the population. Geographic distance and time

influenced whether any two isolates would be the same or different RF, with increases in either making it more likely isolates were independent RFs. In addition, and for reasons that remain unclear, different serotypes exhibited different half-lives, defined as the time between appearance and disappearance of a particular RF [24, 25].

Whilst outbreaks from polio vaccine-derived recombinants has been recognised as relatively frequent and is now monitored, this is not always the case and recombination often has unpredictable outcomes. In the alphaviruses Sindbis and Eastern Equine Encephalitis virus, a recombination event between the structural and non-structural 'modules' led to the emergence of Western Equine Encephalitis virus (WEE), which, although rarely causing symptomatic infections, highlights the importance of recombination in the emergence of novel pathogens [26]. Additionally, recombination between structural genes can also impact virus host range by altering receptor usage, as was the case in the emergence of SARS coronavirus [27] [28].

The increasing use of whole genome sequencing clearly demonstrates the impact of recombination on virus evolution. Combined with experimental systems, recombination is now widely recognised as one of the most important drivers of virus evolution in positive-strand RNA viruses. Traditionally, norovirus classification relied on sequencing of the ORF1 polymerase region only. The rise in identification of naturally occurring recombinants between ORF1 and ORF2 required the inclusion of the capsid region to accurately determine the lineage of each virus [29, 30]. This additional sequencing has uncovered the impact of recombination in the evolution of norovirus genotypes and demonstrated a much higher occurrence of recombination than previously suspected [31-33]. Amongst the three genera of the *Flaviviridae* the incidence of recombination is varied, at least as determined by circulating recombinant forms. While experimental evidence has demonstrated frequent recombination in the pestivirus, bovine viral diarrhoea virus (BVDV), and hepacivirus, hepatitis C virus (HCV) [34-36], additional sequencing and phylogenetic analysis has been required to determine the extent of recombination in other flaviviruses. Recombination events are readily detected in the genomes of the mosquito-borne flaviviruses such as Dengue and Japanese encephalitis virus [37-40], and have been confirmed with limited experimental studies [11, 41]. In contrast, very little evidence of recombination has been observed for West Nile or Yellow Fever virus [40, 42], or any of the tick-borne flaviviruses [43-45], but has been identified when looked for. These differences may be linked to the arthropod vector involved in transmission, but evidence has yet to be presented to support this idea.

Other than the identification of WEE virus as a recombinant virus there has been relatively little research on the role of recombination in alphavirus evolution, or of other members of the *Togaviridae*. The propensity for alphaviruses to recombine has been demonstrated experimentally for Sindbis virus [46, 47] but, with the exception of Chikungunya virus [48, 49], no phylogenetic analyses to identify historical recombination events have been reported. Therefore, although evidence suggests recombination may be more prevalent in some virus families than others, recombination has been found in essentially all virus families where it has been looked for and is undoubtedly a ubiquitous characteristic of single-stranded positive-sense RNA viruses.

Recombination may simply be an accidental by-product of virus replication, reflecting the association and dissociation of RNA template and RdRp or a process, the evolution – and evolutionary retention – of which provides benefit to the individual virus or the virus population. When comparing sexual versus asexual reproduction, an irreversible accumulation of detrimental mutations can severely restrict the evolution of an organism, a phenomenon referred to as Muller's Ratchet [50, 51]. With their error prone polymerases RNA viruses readily mutate, and genomes can accumulate mutations to high levels. While some mutations are beneficial, the accumulation of deleterious mutations leads to virus attenuation [52]. Thus, RNA viruses would greatly benefit from evolving recombination mechanisms to purge these deleterious mutations, while consolidating beneficial ones. Evidence for such requirements has been presented in plant viruses [53] and more recently for poliovirus. In the latter, non-recombinogenic viruses – achieved by polymerase mutagenesis – exhibited a grossly attenuated phenotype and were unable to adapt to certain *in vivo* environments [54, 55]. These studies imply that recombination is a key adaptive process for survival and that the ability to recombine has been evolutionarily selected. However, the apparent absence of recombination in certain genera suggests that either more extensive phylogenetic studies are needed or that, although the process occurs, there is poorly understood functional selection that prevents their fixation in the population. To appreciate this better it is necessary to study the mechanistic process in more detail and to investigate the fate of recombinant genomes.

Defining Recombination

A recombinant virus may be derived following recombination events between two or more individual virus genomes or, as is more rarely observed, recombination between viral and cellular RNA. In the latter process, viral genomes are able to support the insertion of

cellular sequences, for example the incorporation of cellular ubiquitin-like sequences leading to the generation of cytopathogenic strains of BVDV [56, 57]. However, in this article we will focus on the more general process of recombination between virus genomes as recent studies have provided insights into the underlying mechanisms and selection processes that operate.

Many naturally occurring and laboratory-generated recombinants have been characterised from different virus families. To date, two fundamentally distinct mechanisms of recombination have been proposed. It should be noted that, for naturally isolated recombinants, the mechanism of generation cannot be determined as the end products are essentially indistinguishable. For convenience, and reflecting the role of parental genome replication, these mechanisms are referred to as replicative and non-replicative recombination. Perhaps reflecting the historically better-studied process of DNA recombination [58] or a presumption of the underlying mechanism by which they are generated, many studies use the terms homologous and non-homologous to describe the features of the recombinant virus and, more specifically, the recombination junction. Homologous is used to refer to recombinant junctions in which the parental genomes exhibit a wide degree of sequence identity at that position, perhaps not altogether unsurprising in closely related viruses. In contrast, non-homologous recombination refers to junctions, or the process, in which the genetic crossover occurs between poorly conserved regions or unrelated RNA molecules. Additionally, ‘aberrant homologous recombination’ is also used to describe recombination between similar genomes in dissimilar locations. More recently, in light of our studies on the molecular mechanism of recombination ([59], and see below) we introduced the term ‘precise’ to indicate a recombination crossover that exists at the same position in two related parental genomes, and ‘imprecise’ to indicate a junction that maps to two different locations in the parental genomes (whether related or not) (Fig. 1a). This nomenclature has the advantage of not implying anything about the underlying process by which recombinants are generated.

To generate a recombinant virus several principles must be observed. Firstly, a cell must contain two or more viral genomes. Depending on the process these genomes may or may not need to be replication competent (discussed below) but they must be able to physically interact within the cell, and so cannot occupy separate replication complexes or cellular compartments. Secondly, the recombination event must generate a viable genome that is able to replicate and can be packaged into infectious progeny virions. In addition, the resulting virus must be able to survive in competition within a mixed virus population,

although certain transmission routes via a limited inocula, such as in aerosol droplets or by dilution in faecal-contaminated water, may provide sufficient population bottlenecks in which less fit recombinants can proliferate. For recombination between virus genomes of the same species these principles are generally easily met but may explain the lack of recombination observed between viruses of different species and genera.

The *Enterovirus* genera of the *Picornaviridae* family consists of 15 species in total, and while intraspecies recombination is common [17, 60], significant interspecies recombination has not been noted in nature. The assumption here would be that either viruses of different species replicate in separate replication complexes and thus do not have the opportunity to recombine, that recombinant genomes are not viable due to protein-protein or protein-RNA incompatibility, or that recombinants that do arise are insufficiently fit to compete in the environment. The plasticity of the enterovirus 5' NCR has been previously demonstrated through generation of artificial genomes [61, 62] and it is only within this region that any interspecies recombination has been observed, both by *in vitro* generation, and natural isolation, of recombinant viruses [63, 64]. Importantly, by isolating interspecies recombinants, these studies suggest that viruses of different enterovirus species do in fact interact during replication, presumably within the same replication complexes. Interestingly, evidence for inter-family recombination has also been shown with the recent identification of a novel enterovirus circulating in pigs that contains a papain-like protease with high similarity to that of the toroviruses [65-68] although the source and mechanism of acquisition of these sequences remains unclear. The viability and fitness of recombinants is therefore much more likely to explain why recombination is so rarely observed in nature for some viruses.

A key aspect in the generation of viable recombinant viruses is the modular nature of positive-sense single-stranded RNA genomes [9]. Generally, the 5' and 3' termini (non-coding region; NCR) are untranslated and contain sequences and RNA structures implicated in the replication of the genome and the expression of the viral proteins. The remainder of the genome encodes a polyprotein in which the structural and non-structural (respectively, those found in the mature virus particle and those that are produced from the genome but not packaged) coding regions are distinct, but not necessarily separate. Recent metagenomic analysis has demonstrated that the apparent modularity of the genome remains a fundamental characteristic [69]. Whilst this organisation undoubtedly reflects the evolutionary origins of virus families, it also contributes significantly to the process of

recombination by the exchange or acquisition of complete, or near-complete, 'modules' via a single genetic crossover. Mechanistically, how do these crossover events occur?

How do Viruses Recombine?

Replicative Recombination

Genetic recombination in RNA viruses was first identified in the 1960's from the study of poliovirus and, although recombination was readily detected [7, 8, 70, 71], at this time, it was not clear whether recombinants were generated through a DNA-like break-repair mechanism, or through a form of copy-choice, as first proposed by Cooper *et al*. in 1974 [6]. It is now generally accepted that the primary source of recombination is via a replicative process involving a copy-choice mechanism where the viral polymerase switches from one genome to another during negative-strand RNA synthesis [72]. Further progress in determining the mechanisms behind replicative recombination has been slow, in part due to the relatively rare nature of recombination events. The infrequent generation of viable recombinant viruses has meant that, until the widespread use of sequencing, only those able to propagate and become established in a virus population could be isolated or detected. As a result, recombinant genomes are difficult to isolate from the overwhelming majority of parental genomes generated during replication. To overcome the difficulty of isolating recombinant viruses, early studies with poliovirus utilised co-transfection of cells with viral RNAs, each with a different selective genetic marker, such as a temperature sensitive mutation, or resistance to guanidine [6, 7, 72]. Under selective conditions only those genomes that have undergone a recombination event will be permissive for growth and can therefore be isolated and analysed. Similar techniques have also been used to determine the rate of recombination in the coronaviruses [10, 73, 74]. More recently a number of assays have utilised retention of marker genes such as GFP, expressed from the virus genome, to investigate recombination events in poliovirus and flaviviruses [11, 55]. However, the loss of genetic sequences is, semantically, an *in cis* event and may not properly reflect the process of recombination which is, by definition, an *in trans* event involving two genomes.

CRE-REP and Biochemical Assays

To address the need for robust assays to study recombination we have recently developed a cell-based approach known as the CRE-REP assay using poliovirus [59]. The

only viable progeny viruses produced in this assay are recombinants, so facilitating their characterisation. The assay exploits the co-transfection of two compromised RNA genomes that individually cannot generate viable infectious virus, but that will yield viable progeny virus following a successful recombination event (Fig. 1b). The first RNA partner is a poliovirus sub-genomic replicon in which the structural genes have been replaced with a reporter gene, in this case firefly luciferase. Designated the donor, this RNA is fully replication competent but cannot form infectious particles due to the absence of the capsid-coding region. The second RNA partner, the acceptor, contains a cluster of well-characterised point mutations [75, 76] within the *cis*-replicating element (CRE) located within the 2C-coding region. These mutations disrupt the CRE, inhibiting the uridylyation of VPg and consequently preventing positive-strand synthesis [76-78]. Therefore, although this acceptor RNA can be translated and produce negative-strand RNAs it is also unable to generate infectious progeny virus. The designation of donor and acceptor reflects the origin of the polymerase in the resulting recombinants. In single crossover recombinants this is derived from the sub-genomic replicon and occurs from a negative-strand copy-choice recombination event.

In the CRE-REP assay viable recombinants must involve a strand-transfer event located between the functional CRE in the sub-genomic replicon and the junction between the structural and non-structural coding regions of the polyprotein. Hybrids that form outwith these regions would, by definition, lack essential components of the genome. To provide a larger region for recombination we have recently demonstrated that a sub-genomic replicon with the functional CRE re-located to the 3' NCR also functions as a donor template, increases the yield of recombinants and allows recombinants to be isolated throughout the region encoding the non-structural proteins [79]. To facilitate the isolation of early recombinants a cell monolayer permissive for poliovirus replication, but not susceptible to infection, is used. Murine or hamster cells that lack the poliovirus receptor are co-transfected but cannot be subsequently infected by any progeny recombinant viruses. These remain in the supernatant for analysis. Using this assay, we are able to capture recombinants soon after their generation and before they have undergone additional rounds of genome replication.

To our surprise the majority of these early recombinants were found to be greater-than genome length and contained sequence duplications at the site of the recombination junction (Fig. 1b). We designated these imprecise recombinants; since their mechanism of generation was unclear we considered that the term aberrant homologous recombinants –

used in some previous studies to describe similar genomes – which implied that sequence homology was a contributor to their production, was misleading. Where duplications were present, the junctions preferentially straddled the encoded proteolytic cleavage sites in the poliovirus polyprotein. This suggested a selective mechanism whereby recombinants able to encode one non-chimeric copy of a full complement of polyprotein products had a growth advantage. Subsequent studies demonstrated that repeated passaging of the majority of imprecise recombinants results in the deletion of genome duplications and the isolation of recombinants of the correct genome length (which we designated precise recombinants). While this type of imprecise recombinant genome had not been reported in naturally-isolated recombinant enteroviruses, these results were consistent with those subsequently observed from a different poliovirus and coxsackievirus model system [80]. Interestingly, similar recombinants containing sequence duplications have been observed from *in vivo* samples of a BVDV infected animal, seemingly following an initial recombination event with cellular RNA sequences, confirming generation of imprecise recombinants may be a natural step of the recombination process [81]. We proposed that recombination is a biphasic process consisting of an initial, predominantly imprecise, strand-transfer event with selection for replication-competence and the ability to form an infectious virus particle. This is followed by a secondary "resolution" event that – as a consequence of selecting viruses with increased fitness – deletes the genome duplications present in the primary recombinant. Additional evidence for this biphasic recombination process may come from future NGS analysis of recombinant RNA products in co-infected or -transfected cells, or analysis of resolving recombinant virus populations.

The contribution of the viral polymerase to recombination was studied directly using a defined *in vitro* biochemical assay [79]. The RdRp was shown to be necessary and sufficient for catalysing the initial copy-choice template switching event. A growing body of evidence has shown that recombination is intrinsically linked to polymerase activity, with mutations that are known to affect the fidelity of the enzyme directly influencing the rate of recombination between genomes [55, 59, 82, 83]. High fidelity mutants of the poliovirus polymerase, such as a well characterised G64S mutant [84], demonstrate a reduced capacity for recombination compared to wild type levels [59]. These high-fidelity polymerase variants are contenders for inclusion in future non-recombinogenic live attenuated vaccines. In contrast, low fidelity mutants or the addition of sub-lethal levels of ribavirin, an antiviral known to increase the poliovirus polymerase error rate, results in increased yields of recombinant viruses [59, 79]. Mechanistically it remains to be determined how changes in fidelity influence the strand-transfer event. One possibility is that this may be due to

misincorporation leading to polymerase pausing or template dissociation [85]. Although the biochemically-defined *in vitro* recombination assay [79] demonstrates that the polymerase alone is necessary and sufficient to catalyse strand-transfer, there is some evidence that other features of the virus genome contribute to the process of recombination. For example, a chaperone-like activity of the poliovirus 3AB protein may assist in recombination via helix-destabilisation and promotion of RNA annealing between the two parental genomes at the site of the polymerase strand-transfer [86], in a similar manner to the enhancement of recombination by the nucleocapsid protein of HIV [87, 88].

RNA Structure and Sequence Influences on Recombination

Many studies have focused on the characteristics of the junction between parental genomes in naturally or experimentally isolated recombinants, in particular inferring involvement for RNA secondary structures and sequence motifs in the process. For example, the presence of RNA structures, such as hairpin loops, has been linked to the promotion of recombination hotspots in a number of plant viruses within the *Tombusviridae* [89-91]. One suggestion is that hairpin and stem-loop type structures, when present in the donor RNA, will promote the dissociation of the polymerase so initiating a crossover event. Structures present in the acceptor RNA may also influence where the polymerase will re-initiate transcription, with sites upstream of hairpins seemingly favoured to prevent repeated dissociation of the polymerase [89]. Conflicting evidence has been presented however for a different plant virus, Brome mosaic virus (BMV). The tri-partite genome of BMV has been shown to readily undergo both homologous and non-homologous recombination (precise and imprecise respectively using our terminology) between its three RNA components [92, 93]. In one report a specific hairpin structure in RNA3 was found to correlate with the occurrence of non-homologous crossovers in a site-specific manner [94], while an earlier study found no evidence for a role of secondary structures [95]. Similarly, a positive correlation between the presence of RNA structures and recombination hotspots has also been identified for members of the *Flaviviridae* [41], the *Coronaviridae* [96] and poliovirus. In the latter, Runckel *et al*. [97] used NGS to analyse the recombinant progeny between poliovirus genomes bearing numerous translationally-silent genetic tags. In this study they observed increased recombination frequency in regions of localised RNA structure in the largely unstructured genome [98].

However, in analysis of a limited number of recombinant progeny generated in the CRE-REP assay [59] no such link between RNA structure and the recombination junction –

either before or after resolution – could be demonstrated. Furthermore, if the biphasic nature of the process of recombination is generally applicable, the junctions in the chimeric genomes characterised in all previous studies result from fitness selection occurring during the resolution event, rather than recombination *per se*. To address this directly, we have recently investigated recombination between templates engineered to significantly increase or decrease the gross level of RNA structure and show that this did not influence the location of primary or resolved junctions (Bentley *et al*., in preparation).

It is clear that further research is required in this area to fully establish if and how RNA secondary structure influences recombination, and if so, whether this is a universal mechanism of control, or differs between virus families.

Less well studied has been the role of the RNA sequences *per se* in influencing recombination, either as homology (identity) between templates, or as particular sequences that mediate polymerase dissociation and/or re-association. Several studies have reported that the distribution of recombination junctions is biased towards regions of sequence identity between RNA templates [89, 94, 95, 99]. These regions are predicted to be involved in heteroduplex formation between templates, so facilitating polymerase template switching. In agreement with this, the extent of sequence homology is reported to positively influence recombination frequency [95, 100]. However, the majority of these studies pre-date evidence for a biphasic recombination mechanism, so may instead reflect a role for sequence identity in an *in cis* resolution event, a conclusion in agreement with the influence of flanking sequences on reporter gene retention [55]. In our studies of poliovirus, intertypic recombination is less frequent that intratypic, suggesting a positive influence of sequence identity on recombination. However, analysis of intratypic recombinants, where there are regions of limited sequence identity, show no correlation between these regions and the recombination junctions [59, 72].

The nucleotide composition of recombining RNA templates has also been linked to recombination frequencies, with contradictory evidence again presented for viruses of different families. For BMV, AU rich regions were found to be associated with recombination hotspots, while altering the nucleotide composition to create GC rich regions was associated with a suppression of recombination [101-103]. This was also found to be the case for the pestivirus BVDV [104]. In contrast, GC rich regions are reported to be associated with a higher frequency of recombination in poliovirus [97].

It may be that the influence of structure or sequence on recombination varies between viruses and is likely influenced by the methods used to define, generate and analyse

recombinants. Developing techniques and methodologies to investigate precisely how the polymerase interacts with structural and sequence motifs before, and during, the template switching event will be crucial to unravelling the currently mixed evidence of the influence of the RNA template(s) on replicative recombination.

Non-replicative Recombination

Non-replicative recombination was first proposed as a mechanism distinct from that of the copy-choice mediated replicative process from studies of bacteriophage Qß [105]. In these, recombinants are generated *in vitro* – following co-transfection – from overlapping fragments of viral RNA that are individually deficient in their ability to replicate; for example, a 5' partner with the polymerase-encoding gene and the 3' NCR deleted and a 3' partner with deletion of the 5' NCR (Fig. 1c). Those sequences missing from the 5' partner are found in the 3' partner and *vice versa* and, upon co-transfection into cells, these partial genomes can generate viable progeny virus if they can be suitably joined together with a phosphodiester bond. In early studies of this process the inclusion of an intact Qß replicase coding region in the reactions meant that replicative recombination could not be ruled out entirely. However, the identification of recombinants that could only be generated in an end-to-end joining type reaction led to speculation of an alternative mechanism. Subsequently, this mechanism was confirmed for poliovirus, BVDV and HCV [34-36, 106-108]. More recent studies have demonstrated that neither the 5' or 3' components that undergo non-replicative recombination apparently need to be translated, strongly implying that viral proteins are not involved in the process [108].

While non-replicative recombination is arguably a cleaner system with which to generate recombinant viruses for study *in vitro*, the biological significance of this process, and its contribution to recombinant yields *in vivo,* is more difficult to establish. In our own studies, using equivalent amounts of RNA in transfections, we have reported that viable progeny from a proven replicative process are ~25x more frequently generated [59]. More recently, by considerable optimisation of the transfection conditions used, we can generate equivalently high yields of recombinant poliovirus by both replicative and non-replicative mechanisms, in the range of $2x10^4$ PFU/ml per μg of transfected RNA (Bentley, unpublished). This implies that the efficiency or frequency of the two processes may be similar. However, the highly artificial nature of these assays – essentially co-transfection of thousands of copies of *in vitro* synthesised RNA per cell – means that these comparisons are only of relevance under experimental conditions. In reality, only if the 5' and 3' partner

RNAs utilised *in vitro* reflect products naturally generated during virus infection will a non-replicative process be relevant to the evolution of viruses observed in nature. In addition, as the end products of both the replicative and non-replicative mechanisms are potentially the same, the origin of recombinants is impossible to determine and therefore the relative importance of replicative and non-replicative recombination remains unclear.

The mechanisms behind non-replicative recombination have been less well studied than those of replicative recombination. Conflicting evidence has been presented regarding the types of end modification needed on the 5' and 3' partner RNAs in order to facilitate end-to-end joining. Chetverin *et al.* found that the 5' partner RNA required a 3' hydroxyl group for efficient recombination in bacteriophage Qß [105], postulating a mechanism by which, following a covalent interaction of the two RNAs, the 3' hydroxyl group attacks a phosphate group in the sugar backbone of the 3' partner RNA resulting in a ligation of the two RNAs similar to that observed in splicing events. This process required the presence of the Qß RdRp however, suggesting that the majority of recombinants were derived using a replicative process. In a poliovirus system lacking all replicative ability, the opposite result was observed with oxidation of the 3' end of the 5' partner increasing rather than decreasing the number of recombinants [107]. For BVDV, RNA molecules with 3' mono-phosphoryl and 5' hydroxyl ends were found to generate significantly higher yields of recombinants [35]. As these ends represent those generated following endoribonuclease-based cleavage, it was postulated that such cleavage events and a subsequent ligation reaction may provide the basis of a non-replicative recombination mechanism. This would imply a substantial or absolute requirement for host-cell proteins in the non-replicative mechanism, a conclusion supported by the observation that viral proteins are not required for this process [36, 108]. This is in marked contrast to our current understanding of the replicative recombination process, which has an absolute requirement for, and is grossly influenced by, the viral RdRp. This leads to the broader question of the involvement of host proteins in the recombination process, whether replicative or not.

Host Factors in Recombination

RNA viruses rely heavily on a variety of host cell factors for their replication (reviewed in [109] and [110]) but the subversion of host proteins for recombination remains little explored. Host cell factors may be directly required for recombination, as is likely the case for the non-replicative mechanism. In contrast, although it has been demonstrated for replicative recombination that the polymerase alone is sufficient for the strand-transfer

events [79], this does not exclude modulatory roles for either host or other viral proteins. At the forefront of research in this area has been the study of tombusvirus recombination [111]. One advantage of studying tombusviruses is their ability to replicate [112] and recombine in yeast. This has allowed for the large scale genetic screening of host factors that may be involved in recombination [113]. From these screens a number of host genes, many involved in RNA degradation pathways, have been identified that suppress viral recombination. The 5' to 3' exoribonuclease Xrn1 is a key component of the mRNA degradation pathway and has been implicated in recombination, in both yeast and plant model systems, as a regulator of the RNA substrates that are available for recombination [114, 115]. A mechanism was proposed in which viral RNAs are first cleaved by a host endoribonuclease, Ngl2p, to generate 5' and 3' RNA fragments. In the presence of Xrn1 the 3' fragments are rapidly degraded such that they are no longer available to participate in the end-to-end joining mechanism commonly observed for recombination in tombusviruses. This mechanism is presumed to be non-replicative and further genetic screening may enable the identification of the cellular ligase that may also be implicated in the process.

The pathways involved in Xrn1-mediated suppression of recombination may be complex and indirectly regulated by additional host factors. Xrn1 activity was found to be inhibited by pAp, the substrate for yeast (MET22) and plant (AHL, SAL1 and FRY1) nucleotidases. Deletion, or 'knockdown' of the nucleotidase lead to increased substrate inhibition of Xrn1, suppressing viral RNA degradation and thus increasing the observed yield of recombinants [116]. Although the authors of these studies suggest that the assays used measure the results of replicative recombination, the RNA fragments generated by the activity of host endoribonucleases such as Ngl2p [114], or MRP [117], are highly similar to the RNA fragments utilised in cell-based non-replicative recombination assays, and possibly highlight the first stage in the generation of non-replicative recombinants. How such fragments are ligated in the *in vitro* assays remains elusive, but further research into the role of Xrn1, and host endoribonucleases, may help shed light on the as yet undefined mechanism of non-replicative recombination.

Recombination, Defective RNAs, Defective Interfering RNAs and Resolution

A common attribute of virus replication for many RNA viruses, including those with positive strand genomes, is the generation of defective (D), and defective-interfering (DI) RNAs [118-120]. Both lack partial genome sequences – usually regions encoding the structural proteins – whilst retaining the sequences and signals required for replication (Fig.

1a). Defective genomes can therefore replicate, some can be encapsidated into capsids provided *in trans* and those termed 'interfering' can effectively compete with full-length genomes, reducing the yield of viable progeny by usurping replication proteins or capsids. Structurally-similar engineered sub-genomic replicons, in which the capsid protein coding regions are replaced with a reporter gene, have been useful tools to study RNA virus replication and packaging [121, 122], but are generally unable to interfere.

The generation of D-RNAs and DI-RNAs has been far more extensively studied for the negative-strand viruses where it has been shown to occur by several means, including via a copy-choice mechanism [123], thus forging a parallel with the generation of recombinant viruses. In theory, D/DI-RNAs could be generated by an imprecise recombination event leading to a sequence deletion, as opposed to the insertions often observed with recombinants generated via the CRE-REP assay. Mechanistically these are analogous events, with the biased output of the CRE-REP assay simply reflecting the absolute requirement for the production of infectious virus particles. While a link between recombination and the generation of DI-RNAs has been postulated for many years, little evidence has been presented to demonstrate a direct link and determine whether DI-RNA generation occurs intra- or inter molecularly. By using cloned cDNAs of naturally occurring DI-RNAs from the closely related corona- and arteriviruses there is evidence to show that recombination can occur between a DI and its helper virus [124, 125]. Although this does not prove that DI-RNA and recombinant generation utilise the same mechanism, it does suggest that recombination may be a mechanism by which DI-RNA's evolve and are maintained in the virus population. Strong evidence for a link between recombination and DI-RNA generation has been shown for alphaviruses however, using Sindbis virus [83]. Using a mutator strain of the virus Poirier *et al*. demonstrated that an increased polymerase error rate was linked to an increase in the recombination rate and that this led to a higher rate of accumulation of DI-RNA's. Again, while this does not prove that DI-RNA's are generated via the same mechanism as recombinants, the two processes are clearly interlinked. Future research into the mechanisms underlying DI-RNA and recombinant generation may provide interesting insights into how closely related these processes are. Finally, analysis of the process of resolution – at its simplest an *in cis* deletion of duplicated sequences acquired during an imprecise recombination event – may demonstrate it is mechanistically identical to the generation of DI genomes.

The Future of Recombination Research

There is still much to discover about recombination in RNA viruses, both in terms of the mechanism and how it relates to virus evolution. With an improved understanding of the mechanisms we may be able to better predict how and where recombinant viruses will arise and be better prepared for the potential consequences. Although the polymerase has been shown to be sufficient for replicative recombination, a number of aspects of its activity still need to be investigated. A point of particular interest is the relationship between polymerase fidelity, misincorporations, and recombination. Are polymerases that introduce mutations at a higher rate more likely to induce recombination simply due to an increased rate of dissociation and re-association between the template and the enzyme? Alternatively, is there a more intricate relationship in the relative contributions of recombination and mutations in the generation of quasispecies necessary for escaping population bottlenecks? In addition, the polymerases of many RNA viruses oligomerise to form higher order structures [126-128]. Do such structures have a role in recombination by influencing the interaction of the enzyme with the RNA templates? In the latter, the role of localised RNA structures in promoting recombination remains unclear. More extensive, longer range or higher order RNA structures, such as genome cyclization [129] or genome-scale ordered RNA structure [98], may have as yet unidentified importance in recombination or resolution.

With the exception of an essential contribution of the RdRp to replicative recombination, roles for additional viral or cellular proteins in the process remain unclear. If requirements do differ between related viruses it may explain the apparent restrictions in recombination observed. Specificity determined by intracellular components could also extend to include the compartmentalisation of the virus genome in replication complexes (logically only in replicative recombination). We and others have demonstrated that recombination is reduced by nocodazole [59, 130], an inhibitor of microtubule polymerisation and consequent coalescence of replication complexes. Even two closely related viruses, occupying replication complexes derived from different membrane compartments, may never get the opportunity to meet and recombine.

As with other areas of virology, the ability to interrogate the entire RNA population using deep sequencing strategies will provide important insights into recombination and resolution. Is the initial recombination (crossover) event truly promiscuous, with no sequence specificity and a random distribution throughout both parental genomes? Similarly, what determines the intermediates and the final products of resolution? Is it solely the functionality and fitness of the resulting virus? A better understanding of these processes will help us comprehend this fundamental evolutionary mechanism. Additionally, by

understanding what leads to the selection of particular recombinants over the countless number that could be generated, we will gain important insights into the functionality of the encoded viral proteins, their interactions with each other and with those of the host.

References

1. **Domingo E, Holland JJ**. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* 1997;51:151-178.
2. **Essere B, Yver M, Gavazzi C, Terrier O, Isel C et al.** Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. *Proc Natl Acad Sci U S A* 2013;110(40):E3840-3848.
3. **McDonald SM, Nelson MI, Turner PE, Patton JT**. Reassortment in segmented RNA viruses: mechanisms and outcomes. *Nat Rev Microbiol* 2016;14(7):448-460.
4. **Muller HJ**. Some Genetic Aspects of Sex. *The American Naturalist* 1932;66(703):118-138.
5. **Han GZ, Worobey M**. Homologous recombination in negative sense RNA viruses. *Viruses* 2011;3(8):1358-1373.
6. **Copper PD, Steiner-Pryor A, Scotti PD, Delong D**. On the nature of poliovirus genetic recombinants. *J Gen Virol* 1974;23(1):41-49.
7. **Sergiescu D, Aubert-Combiescu A, Crainic R**. Recombination between guanidine-resistant and dextran sulfate-resistant mutants of type 1 poliovirus. *J Virol* 1969;3(3):326-330.
8. **LEDINKO N**. Genetic recombination with poliovirus type 1. Studies of crosses between a normal horse serum-resistant mutant and several guanidine-resistant mutants of the same strain. *Virology* 1963;20:107-119.
9. **Simmonds P, Welch J**. Frequency and dynamics of recombination within different species of human enteroviruses. *J Virol* 2006;80(1):483-493.
10. **Makino S, Keck JG, Stohlman SA, Lai MM**. High-frequency RNA recombination of murine coronaviruses. *J Virol* 1986;57(3):729-737.
11. **Taucher C, Berger A, Mandl CW**. A trans-complementing recombination trap demonstrates a low propensity of flaviviruses for intermolecular recombination. *J Virol* 2010;84(1):599-611.
12. **Mindich L, Qiao X, Onodera S, Gottlieb P, Strassman J**. Heterologous recombination in the double-stranded RNA bacteriophage phi 6. *J Virol* 1992;66(5):2605-2610.
13. **Lai MM**. RNA recombination in animal and plant viruses. *Microbiol Rev* 1992;56(1):61-79.
14. **Bujarski JJ**. Genetic recombination in plant-infecting messenger-sense RNA viruses: overview and research perspectives. *Front Plant Sci* 2013;4:68.
15. **Cammack N, Phillips A, Dunn G, Patel V, Minor PD**. Intertypic genomic rearrangements of poliovirus strains in vaccinees. *Virology* 1988;167(2):507-514.
16. **Cuervo NS, Guillot S, Romanenkova N, Combiescu M, Aubert-Combiescu A et al.** Genomic features of intertypic recombinant sabin poliovirus strains excreted by primary vaccinees. *J Virol* 2001;75(13):5740-5751.
17. **Kew O, Morris-Glasgow V, Landaverde M, Burns C, Shaw J et al.** Outbreak of poliomyelitis in Hispaniola associated with circulating type 1 vaccine-derived poliovirus. *Science* 2002;296(5566):356-359.
18. **Rakoto-Andrianarivelo M, Gumede N, Jegouic S, Balanant J, Andriamamonjy SN et al.** Reemergence of recombinant vaccine-derived poliovirus outbreak in Madagascar. *J Infect Dis* 2008;197(10):1427-1435.
19. **Liang X, Zhang Y, Xu W, Wen N, Zuo S et al.** An outbreak of poliomyelitis caused by type 1 vaccine-derived poliovirus in China. *J Infect Dis* 2006;194(5):545-551.

20. **Joffret ML, Jégouic S, Bessaud M, Balanant J, Tran C et al.** Common and diverse features of cocirculating type 2 and 3 recombinant vaccine-derived polioviruses isolated from patients with poliomyelitis and healthy children. *J Infect Dis* 2012;205(9):1363-1373.
21. **Lukashev AN, Lashkevich VA, Ivanova OE, Koroleva GA, Hinkkanen AE et al.** Recombination in circulating enteroviruses. *J Virol* 2003;77(19):10423-10431.
22. **Oberste MS, Peñaranda S, Pallansch MA**. RNA recombination plays a major role in genomic change during circulation of coxsackie B viruses. *J Virol* 2004;78(6):2948-2955.
23. **Oberste MS, Maher K, Pallansch MA**. Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. *J Virol* 2004;78(2):855-867.
24. **McWilliam Leitch EC, Cabrerizo M, Cardosa J, Harvala H, Ivanova OE et al.** Evolutionary dynamics and temporal/geographical correlates of recombination in the human enterovirus echovirus types 9, 11, and 30. *J Virol* 2010;84(18):9292-9300.
25. **McWilliam Leitch EC, Cabrerizo M, Cardosa J, Harvala H, Ivanova OE et al.** The association of recombination events in the founding and emergence of subgenogroup evolutionary lineages of human enterovirus 71. *J Virol* 2012;86(5):2676-2685.
26. **Hahn CS, Lustig S, Strauss EG, Strauss JH**. Western equine encephalitis virus is a recombinant virus. *Proc Natl Acad Sci U S A* 1988;85(16):5997-6001.
27. **Ge XY, Li JL, Yang XL, Chmura AA, Zhu G et al.** Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 2013;503(7477):535-538.
28. **Graham RL, Baric RS**. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol* 2010;84(7):3134-3146.
29. **Bull RA, Tanaka MM, White PA**. Norovirus recombination. *J Gen Virol* 2007;88(Pt 12):3347-3359.
30. **Bull RA, Hansman GS, Clancy LE, Tanaka MM, Rawlinson WD et al.** Norovirus recombination in ORF1/ORF2 overlap. *Emerg Infect Dis* 2005;11(7):1079-1085.
31. **Mahar JE, Bok K, Green KY, Kirkwood CD**. The importance of intergenic recombination in norovirus GII.3 evolution. *J Virol* 2013;87(7):3687-3698.
32. **Cannon JL, Barclay L, Collins NR, Wikswo ME, Castro CJ et al.** Genetic and Epidemiologic Trends of Norovirus Outbreaks in the United States from 2013 to 2016 Demonstrated Emergence of Novel GII.4 Recombinant Viruses. *J Clin Microbiol* 2017;55(7):2208-2221.
33. **Eden JS, Tanaka MM, Boni MF, Rawlinson WD, White PA**. Recombination within the pandemic norovirus GII.4 lineage. *J Virol* 2013;87(11):6270-6282.
34. **Gallei A, Pankraz A, Thiel HJ, Becher P**. RNA recombination in vivo in the absence of viral replication. *J Virol* 2004;78(12):6271-6281.
35. **Austermann-Busch S, Becher P**. RNA structural elements determine frequency and sites of nonhomologous recombination in an animal plus-strand RNA virus. *J Virol* 2012;86(13):7393-7402.
36. **Scheel TK, Galli A, Li YP, Mikkelsen LS, Gottwein JM et al.** Productive homologous and non-homologous recombination of hepatitis C virus in cell culture. *PLoS Pathog* 2013;9(3):e1003228.
37. **Holmes EC, Worobey M, Rambaut A**. Phylogenetic evidence for recombination in dengue virus. *Mol Biol Evol* 1999;16(3):405-409.
38. **Worobey M, Rambaut A, Holmes EC**. Widespread intra-serotype recombination in natural populations of dengue virus. *Proc Natl Acad Sci U S A* 1999;96(13):7352-7357.
39. **Carney J, Daly JM, Nisalak A, Solomon T**. Recombination and positive selection identified in complete genome sequences of Japanese encephalitis virus. *Arch Virol* 2012;157(1):75-83.
40. **Twiddy SS, Holmes EC**. The extent of homologous recombination in members of the genus Flavivirus. *J Gen Virol* 2003;84(Pt 2):429-440.
41. **Chuang CK, Chen WJ**. Experimental evidence that RNA recombination occurs in the Japanese encephalitis virus. *Virology* 2009;394(2):286-297.

42. **Pickett BE, Lefkowitz EJ**. Recombination in West Nile Virus: minimal contribution to genomic diversity. *Virol J* 2009;6:165.
43. **Bertrand Y, Töpel M, Elväng A, Melik W, Johansson M**. First dating of a recombination event in mammalian tick-borne flaviviruses. *PLoS One* 2012;7(2):e31981.
44. **Bertrand YJ, Johansson M, Norberg P**. Revisiting Recombination Signal in the Tick-Borne Encephalitis Virus: A Simulation Approach. *PLoS One* 2016;11(10):e0164435.
45. **Norberg P, Roth A, Bergström T**. Genetic recombination of tick-borne flaviviruses among wild-type strains. *Virology* 2013;440(2):105-116.
46. **Weiss BG, Schlesinger S**. Recombination between Sindbis virus RNAs. *J Virol* 1991;65(8):4017-4025.
47. **Hill KR, Hajjou M, Hu JY, Raju R**. RNA-RNA recombination in Sindbis virus: roles of the 3' conserved motif, poly(A) tail, and nonviral sequences of template RNAs in polymerase recognition and template switching. *J Virol* 1997;71(4):2693-2704.
48. **Casal PE, Chouhy D, Bolatti EM, Perez GR, Stella EJ et al.** Evidence for homologous recombination in Chikungunya Virus. *Mol Phylogenet Evol* 2015;85:68-75.
49. **Cui J, Gao M, Ren X**. Phylogeny and homologous recombination in Chikungunya viruses. *Infect Genet Evol* 2011;11(8):1957-1963.
50. **Felsenstein J**. The evolutionary advantage of recombination. *Genetics* 1974;78(2):737-756.
51. **Muller HJ**. The relation of recombination to mutational advance. *Mutat Res* 1964;106:2-9.
52. **Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R**. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 2006;439(7074):344-348.
53. **Allison R, Thompson C, Ahlquist P**. Regeneration of a functional RNA virus genome by recombination between deletion mutants and requirement for cowpea chlorotic mottle virus 3a and coat genes for systemic infection. *Proc Natl Acad Sci U S A* 1990;87(5):1820-1824.
54. **Xiao Y, Dolan PT, Goldstein EF, Li M, Farkov M et al.** Poliovirus intrahost evolution is required to overcome tissue-specific innate immune responses. *Nat Commun* 2017;8(1):375.
55. **Xiao Y, Rouzine IM, Bianco S, Acevedo A, Goldstein EF et al.** RNA Recombination Enhances Adaptability and Is Required for Virus Spread and Virulence. *Cell Host Microbe* 2016;19(4):493-503.
56. **Meyers G, Tautz N, Dubovi EJ, Thiel HJ**. Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. *Virology* 1991;180(2):602-616.
57. **Becher P, Tautz N**. RNA recombination in pestiviruses: cellular RNA sequences in viral genomes highlight the role of host factors for viral persistence and lethal disease. *RNA Biol* 2011;8(2):216-224.
58. **Holliday R**. A mechanism for gene conversion in fungi. *Genetical Research* 1964;5(2):282-304.
59. **Lowry K, Woodman A, Cook J, Evans DJ**. Recombination in enteroviruses is a biphasic process involving the generation of greater-than genome length 'imprecise' intermediates. *PLoS Pathogens* 2013;Submitted.
60. **Hughes PJ, North C, Minor PD, Stanway G**. The complete nucleotide sequence of coxsackievirus A21. *J Gen Virol* 1989;70 (Pt 11):2943-2952.
61. **Rohll JB, Percy N, Ley R, Evans DJ, Almond JW et al.** The 5'-untranslated regions of picornavirus RNAs contain independent functional domains essential for RNA replication and translation. *J Virol* 1994;68(7):4384-4391.
62. **Schibler M, Gerlach D, Martinez Y, Belle SV, Turin L et al.** Experimental human rhinovirus and enterovirus interspecies recombination. *J Gen Virol* 2012;93(Pt 1):93-101.
63. **Muslin C, Joffret ML, Pelletier I, Blondel B, Delpeyroux F**. Evolution and Emergence of Enteroviruses through Intra- and Inter-species Recombination: Plasticity and Phenotypic Impact of Modular Genetic Exchanges in the 5' Untranslated Region. *PLoS Pathog* 2015;11(11):e1005266.

64. **Yozwiak NL, Skewes-Cox P, Gordon A, Saborio S, Kuan G et al.** Human enterovirus 109: a novel interspecies recombinant enterovirus isolated from a case of acute pediatric respiratory illness in Nicaragua. *J Virol* 2010;84(18):9047-9058.
65. **Shang P, Misra S, Hause B, Fang Y**. A Naturally Occurring Recombinant Enterovirus Expresses a Torovirus Deubiquitinase. *J Virol* 2017;91(14).
66. **Knutson TP, Velayudhan BT, Marthaler DG**. A porcine enterovirus G associated with enteric disease contains a novel papain-like cysteine protease. *J Gen Virol* 2017;98(6):1305-1310.
67. **Conceição-Neto N, Theuns S, Cui T, Zeller M, Yinda CK et al.** Identification of an enterovirus recombinant with a torovirus-like gene insertion during a diarrhea outbreak in fattening pigs. *Virus Evol* 2017;3(2):vex024.
68. **Wang Y, Zhang W, Liu Z, Fu X, Yuan J et al.** Full-length and defective enterovirus G genomes with distinct torovirus protease insertions are highly prevalent on a Chinese pig farm. *Arch Virol* 2018.
69. **Shi M, Lin XD, Chen X, Tian JH, Chen LJ et al.** The evolutionary history of vertebrate RNA viruses. *Nature* 2018;556(7700):197-202.
70. **Cooper PD**. A genetic map of poliovirus temperature-sensitive mutants. *Virology* 1968;35(4):584-596.
71. **HIRST GK**. Genetic recombination with Newcastle disease virus, polioviruses, and influenza. *Cold Spring Harb Symp Quant Biol* 1962;27:303-309.
72. **Kirkegaard K, Baltimore D**. The mechanism of RNA recombination in poliovirus. *Cell* 1986;47(3):433-443.
73. **Baric RS, Fu K, Schaad MC, Stohlman SA**. Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. *Virology* 1990;177(2):646-656.
74. **Fu K, Baric RS**. Map locations of mouse hepatitis virus temperature-sensitive mutants: confirmation of variable rates of recombination. *J Virol* 1994;68(11):7458-7466.
75. **Goodfellow IG, Kerrigan D, Evans DJ**. Structure and function analysis of the poliovirus cis-acting replication element (CRE). *RNA* 2003;9(1):124-137.
76. **Rieder E, Paul AV, Kim DW, van Boom JH, Wimmer E**. Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation. *J Virol* 2000;74(22):10371-10380.
77. **Paul AV, Rieder E, Kim DW, van Boom JH, Wimmer E**. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. *J Virol* 2000;74(22):10359-10370.
78. **Goodfellow I, Polacek C, Andino R, Evans D**. The poliovirus 2C cis-acting replication element-mediated uridylylation of VPg is not required for synthesis of negative-sense genomes. *J Gen Virol* 2003;84(Pt 9):2359-2363.
79. **Woodman A, Arnold JJ, Cameron CE, Evans DJ**. Biochemical and genetic analysis of the role of the viral polymerase in enterovirus recombination. *Nucleic Acids Res* 2016;44(14):6883-6895.
80. **Holmblat B, Jégouic S, Muslin C, Blondel B, Joffret ML et al.** Nonhomologous recombination between defective poliovirus and coxsackievirus genomes suggests a new model of genetic plasticity for picornaviruses. *MBio* 2014;5(4):e01119-01114.
81. **Fricke J, Gunn M, Meyers G**. A family of closely related bovine viral diarrhea virus recombinants identified in an animal suffering from mucosal disease: new insights into the development of a lethal disease in cattle. *Virology* 2001;291(1):77-90.
82. **Kempf BJ, Peersen OB, Barton DJ**. Poliovirus Polymerase Leu420 Facilitates RNA Recombination and Ribavirin Resistance. *J Virol* 2016;90(19):8410-8421.
83. **Poirier EZ, Mounce BC, Rozen-Gagnon K, Hooikaas PJ, Stapleford KA et al.** Low-fidelity polymerases of alphaviruses recombine at higher rates to overproduce defective interfering particles. *J Virol* 2015.

84. **Pfeiffer JK, Kirkegaard K**. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog* 2005;1(2):e11.

85. **Dulin D, Vilfan ID, Berghuis BA, Hage S, Bamford DH et al.** Elongation-Competent Pauses Govern the Fidelity of a Viral RNA-Dependent RNA Polymerase. *Cell Rep* 2015.

86. **DeStefano JJ, Titilope O**. Poliovirus protein 3AB displays nucleic acid chaperone and helix-destabilizing activities. *J Virol* 2006;80(4):1662-1671.

87. **Negroni M, Buc H**. Copy-choice recombination by reverse transcriptases: reshuffling of genetic markers mediated by RNA chaperones. *Proc Natl Acad Sci U S A* 2000;97(12):6385-6390.

88. **Negroni M, Buc H**. Recombination during reverse transcription: an evaluation of the role of the nucleocapsid protein. *J Mol Biol* 1999;286(1):15-31.

89. **White KA, Morris TJ**. RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. *RNA* 1995;1(10):1029-1040.

90. **Nagy PD, Zhang C, Simon AE**. Dissecting RNA recombination in vitro: role of RNA sequences and the viral replicase. *EMBO J* 1998;17(8):2392-2403.

91. **Carpenter CD, Oh JW, Zhang C, Simon AE**. Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. *J Mol Biol* 1995;245(5):608-622.

92. **Bruyere A, Wantroba M, Flasinski S, Dzianott A, Bujarski JJ**. Frequent homologous recombination events between molecules of one RNA component in a multipartite RNA virus. *J Virol* 2000;74(9):4214-4219.

93. **Bujarski JJ, Kaesberg P**. Genetic recombination between RNA components of a multipartite plant virus. *Nature* 1986;321(6069):528-531.

94. **Figlerowicz M**. Role of RNA structure in non-homologous recombination between genomic molecules of brome mosaic virus. *Nucleic Acids Res* 2000;28(8):1714-1723.

95. **Nagy PD, Bujarski JJ**. Efficient system of homologous RNA recombination in brome mosaic virus: sequence and structure requirements and accuracy of crossovers. *J Virol* 1995;69(1):131-140.

96. **Rowe CL, Fleming JO, Nathan MJ, Sgro JY, Palmenberg AC et al.** Generation of coronavirus spike deletion variants by high-frequency recombination at regions of predicted RNA secondary structure. *J Virol* 1997;71(8):6183-6190.

97. **Runckel C, Westesson O, Andino R, DeRisi JL**. Identification and manipulation of the molecular determinants influencing poliovirus recombination. *PLoS Pathog* 2013;9(2):e1003164.

98. **Simmonds P, Tuplin A, Evans DJ**. Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: Implications for virus evolution and host persistence. *RNA* 2004;10(9):1337-1351.

99. **Cascone PJ, Haydar TF, Simon AE**. Sequences and structures required for recombination between virus-associated RNAs. *Science* 1993;260(5109):801-805.

100. **Draghici HK, Varrelmann M**. Evidence for similarity-assisted recombination and predicted stem-loop structure determinant in potato virus X RNA recombination. *J Gen Virol* 2010;91(Pt 2):552-562.

101. **Shapka N, Nagy PD**. The AU-rich RNA recombination hot spot sequence of Brome mosaic virus is functional in tombusviruses: implications for the mechanism of RNA recombination. *J Virol* 2004;78(5):2288-2300.

102. **Nagy PD, Bujarski JJ**. Engineering of homologous recombination hotspots with AU-rich sequences in brome mosaic virus. *J Virol* 1997;71(5):3799-3810.

103. **Nagy PD, Bujarski JJ**. Silencing homologous RNA recombination hot spots with GC-rich sequences in brome mosaic virus. *J Virol* 1998;72(2):1122-1130.

104. **Kim MJ, Kao C**. Factors regulating template switch in vitro by viral RNA-dependent RNA polymerases: implications for RNA-RNA recombination. *Proc Natl Acad Sci U S A* 2001;98(9):4972-4977.

105. **Chetverin AB, Chetverina HV, Demidenko AA, Ugarov VI**. Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. *Cell* 1997;88(4):503-513.
106. **Gmyl AP, Korshenko SA, Belousov EV, Khitrina EV, Agol VI**. Nonreplicative homologous RNA recombination: promiscuous joining of RNA pieces? *RNA* 2003;9(10):1221-1231.
107. **Gmyl AP, Belousov EV, Maslova SV, Khitrina EV, Chetverin AB et al.** Nonreplicative RNA recombination in poliovirus. *J Virol* 1999;73(11):8958-8965.
108. **Kleine Büning M, Meyer D, Austermann-Busch S, Roman-Sosa G, Rümenapf T et al.** Nonreplicative RNA Recombination of an Animal Plus-Strand RNA Virus in the Absence of Efficient Translation of Viral Proteins. *Genome Biol Evol* 2017;9(4):817-829.
109. **Ahlquist P, Noueiry AO, Lee WM, Kushner DB, Dye BT**. Host factors in positive-strand RNA virus genome replication. *J Virol* 2003;77(15):8181-8186.
110. **Nagy PD, Pogany J**. The dependence of viral RNA replication on co-opted host factors. *Nat Rev Microbiol* 2011;10(2):137-149.
111. **Nagy PD**. The roles of host factors in tombusvirus RNA recombination. *Adv Virus Res* 2011;81:63-84.
112. **Ishikawa M, Janda M, Krol MA, Ahlquist P**. In vivo DNA expression of functional brome mosaic virus RNA replicons in Saccharomyces cerevisiae. *J Virol* 1997;71(10):7781-7790.
113. **Serviene E, Shapka N, Cheng CP, Panavas T, Phuangrat B et al.** Genome-wide screen identifies host genes affecting viral RNA recombination. *Proc Natl Acad Sci U S A* 2005;102(30):10545-10550.
114. **Cheng CP, Serviene E, Nagy PD**. Suppression of viral RNA recombination by a host exoribonuclease. *J Virol* 2006;80(6):2631-2640.
115. **Jaag HM, Nagy PD**. Silencing of Nicotiana benthamiana Xrn4p exoribonuclease promotes tombusvirus RNA accumulation and recombination. *Virology* 2009;386(2):344-352.
116. **Jaag HM, Nagy PD**. The combined effect of environmental and host factors on the emergence of viral RNA recombinants. *PLoS Pathog* 2010;6(10):e1001156.
117. **Jaag HM, Lu Q, Schmitt ME, Nagy PD**. Role of RNase MRP in viral RNA degradation and RNA recombination. *J Virol* 2011;85(1):243-253.
118. **Kajigaya S, Arakawa H, Kuge S, Koi T, Imura N et al.** Isolation and characterization of defective-interfering particles of poliovirus Sabin 1 strain. *Virology* 1985;142(2):307-316.
119. **Pénzes Z, Wroe C, Brown TD, Britton P, Cavanagh D**. Replication and packaging of coronavirus infectious bronchitis virus defective RNAs lacking a long open reading frame. *J Virol* 1996;70(12):8660-8668.
120. **Brinton MA**. Analysis of extracellular West Nile virus particles produced by cell cultures from genetically resistant and susceptible mice indicates enhanced amplification of defective interfering particles by resistant cultures. *J Virol* 1983;46(3):860-870.
121. **Barclay W, Li Q, Hutchinson G, Moon D, Richardson A et al.** Encapsidation studies of poliovirus subgenomic replicons. *J Gen Virol* 1998;79 (Pt 7):1725-1734.
122. **Lohmann V, Körner F, Koch J, Herian U, Theilmann L et al.** Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* 1999;285(5424):110-113.
123. **Lazzarini RA, Keene JD, Schubert M**. The origins of defective interfering particles of the negative-strand RNA viruses. *Cell* 1981;26(2 Pt 2):145-154.
124. **Furuya T, Macnaughton TB, La Monica N, Lai MM**. Natural evolution of coronavirus defective-interfering RNA involves RNA recombination. *Virology* 1993;194(1):408-413.
125. **Molenkamp R, Greve S, Spaan WJ, Snijder EJ**. Efficient homologous RNA recombination and requirement for an open reading frame during replication of equine arteritis virus defective interfering RNAs. *J Virol* 2000;74(19):9062-9070.

126. **Wang QM, Hockman MA, Staschke K, Johnson RB, Case KA et al.** Oligomerization and cooperative RNA synthesis activity of hepatitis C virus RNA-dependent RNA polymerase. *J Virol* 2002;76(8):3865-3872.
127. **Hansen JL, Long AM, Schultz SC**. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* 1997;5(8):1109-1122.
128. **Kaiser WJ, Chaudhry Y, Sosnovtsev SV, Goodfellow IG**. Analysis of protein-protein interactions in the feline calicivirus replication complex. *J Gen Virol* 2006;87(Pt 2):363-368.
129. **Villordo SM, Gamarnik AV**. Genome cyclization as strategy for flavivirus RNA replication. *Virus Res* 2009;139(2):230-239.
130. **Egger D, Bienz K**. Intracellular location and translocation of silent and active poliovirus replication complexes. *J Gen Virol* 2005;86(Pt 3):707-718.

Figure Legend

Fig.1. Recombinant single-stranded positive-sense RNA viruses, their nomenclature and generation. Schematic genomes are representative of poliovirus, the prototype enterovirus, with numbering of the proteins within by the polyprotein; 1-4 are the structural proteins (VP4, 2, 3 and 1), 5-11 are the non-structural proteins ($2A^{pro}$, 2B, 2C, 3A, 3B[VPg], $3C^{pro}$, $3D^{pol}$), with primed numbers indicating partial or incomplete protein coding regions. **(a)** Co-infection of cells with two viruses (represented by Genome 1 and Genome 2) may result in a number of replication-competent genomes. Precise recombinants are parental-length with a junction and no extraneous sequences. Imprecise recombinants contain duplications at the crossover region. For convenience we have omitted the potential reciprocal products with Genome 2 forming the 5' end of the progeny. Defective RNA genomes contain an *in cis* deletion, typically within the capsid-coding region. Single crossover events are only shown, though we and others have detected double crossovers in natural and experimentally generated recombinants. **(b)** The CRE-REP assay. The Acceptor genome (blue) bears a well-defined modification of the essential *cis*-replicating element (CRE) in the 2C coding region that prevents positive-strand synthesis. The Donor genome (red; so-called because the polymerase in the progeny is derived from this parental genome) has the capsid-coding sequences replaced with a reporter gene (Luc). Donor and Acceptor RNA are transfected into permissive cells and undergo recombination during negative strand synthesis to generate imprecise recombinants. Subsequent rounds of replication, within the original cell or upon serial passage in uninfected cells, leads to resolution resulting in the loss of the genome duplications within the imprecise recombinant. **(c)** Non-replicative recombination assay. Donor (red) and Acceptor (blue) partner RNAs are generated *in vitro* by truncation of the 3' or 5' ends, respectively of the parental virus genome.

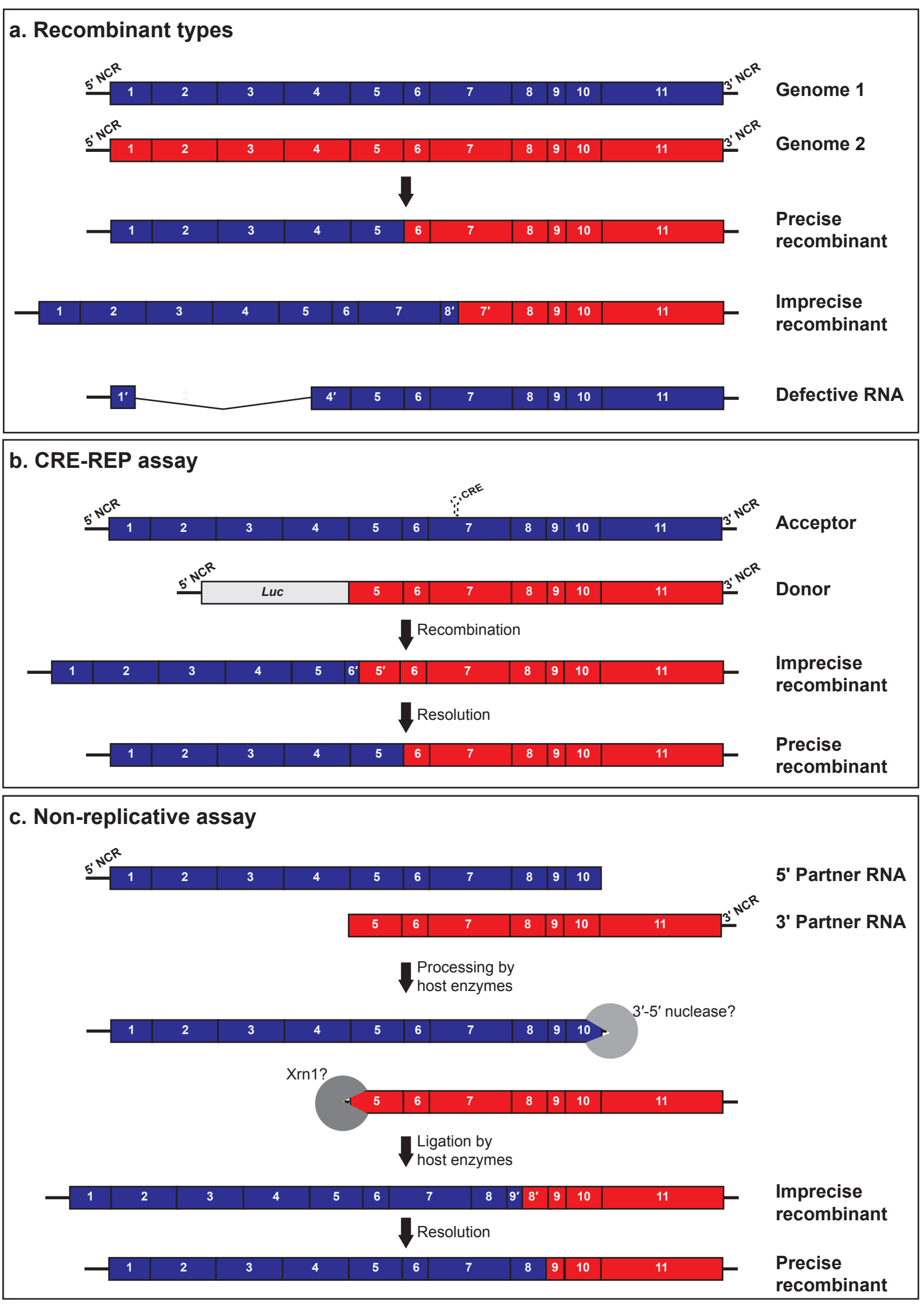
Co-transfected RNAs are presumed to be processed by host endo- and exonucleases and subsequently ligated by host RNA ligases. The initial product, an imprecise recombinant – if replication-competent – can undergo subsequent resolution in which duplicated sequences within the genome are lost.

Acknowledgements:

The authors thank Craig Cameron, Raul Andino and Peter Simmonds for stimulating discussions on virus recombination.

Funding: KB is supported by BBSRC award BB/M009343/1 to DJE

Conflicts of interest: The authors declare that there are no conflicts of interest


a. Recombinant types
5′ NCR
3′ NCR
Genome 1
Genome 2
Precise recombinant
Imprecise recombinant
Defective RNA
b. CRE-REP assay
CRE
Acceptor
Luc
Donor
Recombination
Resolution
c. Non-replicative assay
5' Partner RNA
3' Partner RNA
Processing by host enzymes
3′-5′ nuclease?
Xrn1?
Ligation by host enzymes